

From “Oh, OK” to “Ah, yes” to “Aha!”: Hyper-systemizing and the rewards of insight

Chris Fields

Apdo 363-4013
Atenas 20501
Costa Rica

chrisfields38@gmail.com

Abstract

Hyper-systemizers are individuals displaying an unusually strong bias toward systemizing, i.e. toward explaining events and solving problems by appeal to mechanisms that do not involve intentions or agency. Hyper-systemizing in combination with deficit mentalizing capability typically presents clinically as an autism spectrum disorder; however, the development of hyper-systemizing in combination with normal-range mentalizing capability is not well characterized. Based on a review and synthesis of clinical, observational, experimental, and neurofunctional studies, it is hypothesized that repeated episodes of insightful problem solving by systemizing result in attentional and motivational sensitization toward further systemizing via progressive and chronic deactivation of the default network. This hypothesis is distinguished from alternatives, and its correlational and causal implications are discussed. Predictions of the default-deactivation model accessible to survey-based instruments, standard cognitive measures and neurofunctional methods are outlined, and evidence pertaining to them considered.

Keywords: Systemizing, Mentalizing, Development, Insight, Default network, Sensitization, Analogy, ASD

1. Introduction

Scientists, technologists, engineers and mathematicians rely heavily on a problem-solving and explanatory strategy or orientation, termed “systemizing” by Baron-Cohen (2002, 2008), that is characterized by appeals to natural laws, physical mechanisms, algorithms, formal inference rules, or other concepts of causation that do not involve autonomous agency or intentions. Systemizing or “mechanizing” (Crespi and Badcock, 2008) solutions and explanations are explicitly distinguished from “empathizing” or “mentalizing” solutions and explanations, which do appeal to intentional,

autonomous agency and to actions guided by beliefs, desires, goals, fears, worries and other “folk psychological” attributes associated with agency by a theory of mind (ToM) system (Frith and Frith, 1999; 2003). Nearly everyone reasons by systemizing at least some of the time, appealing to gravity instead of desire for union with the Earth to explain the fall of an apple, and to school-room germs not curses hurled by malicious neighbors to explain a child's bout with the flu. Some people, however, spend much or most of their time systemizing, and such “hyper-systemizers” often find roles in society as scientists, technologists, engineers or mathematicians (Baron-Cohen, 2008).

Hyper-systemizers tend to display a collection of personality characteristics typified in life by Albert Einstein and parodied in popular culture by the absent-minded professor or the “geeky” engineer: obsessive focus on abstruse and technical interests, social awkwardness, isolation, aloofness, and disregard for ordinary conventions regarding public appearance and behavior. Empirical studies bear this out; Feist and Gorman (1998) summarize three decades of research on the personality characteristics of scientists as indicating that “(I)f scientists in general are more aloof, asocial and introverted than non-scientists, then these characteristics appear to be even more salient for the scientific elite” (p. 28; “scientists” here includes technologists). Some regard the expression of these characteristics by Einstein and many other eminent hyper-systemizers as clinically significant. Fitzgerald and O'Brien (2007), for example, have retrospectively diagnosed Einstein, Sir Isaac Newton, mathematicians David Hilbert and Kurt Gödel, information scientists Charles Babbage and Norbert Wiener, biologists Charles Darwin and Gregor Mendel, and psychologists John Watson and Alfred Kinsey as sufferers of Asperger's syndrome. While such diagnoses are not universally accepted, they indicate the extent to which a problem-solving strategy can be associated, in some of its most capable employers, with personality traits sometimes regarded as dysfunctional.

As Feist and Gorman (1998) point out, the causal relationship between ability in scientific thinking and the personality characteristics typical of scientists is far from clear, but almost certainly involves an interaction between pre-existing personality and the cognitive demands of scientific work. Based on a review and synthesis of relevant literatures, the present paper proposes that pre-existing personality and cognitive demands interact to progressively sensitize the attentional and motivational systems of some individuals toward systemizing and away from mentalizing, resulting in hyper-systemizers who are not deficit in mentalizing capability, but rather relatively insensitive to cues that ordinarily induce mentalizing. The paper is organized as follows. Section 2 reviews the phenomenology of systemizing and mentalizing as human problem-solving strategies, and explicitly distinguishes an orientation toward one or the other strategy as measurable by survey instruments (Baron-Cohen *et al.*, 2003) from capability with either strategy determined by performance on relevant tasks. The correlations between hyper-systemizing and Autism Spectrum Disorders (ASD) are then reviewed, and the question of whether hyper-systemizing itself is pathological is considered. Section 3 reviews the affective experiences typically accompanying problem-solving by systemizing or mentalizing. The hedonic spectrum from the “Oh, OK” of everyday solution-finding through the “Ah, yes” of mild surprise to the “Aha!” of insight is characterized in terms of increasing intensity of affective reward for current-state to goal-state conflict resolution. Section 4 reviews evidence bearing on the neurocognitive mechanisms that implement the hedonic response to insight, and compares them to the mechanisms proposed by the incentive-sensitization model of drug addiction (Robinson and Berridge, 1993; 2008). It suggests that chronic sensitization to cues associated with systemizing may drive the development of hyper-systemizing. Section 5 reviews evidence indicating that decreased default-network activity is not only a correlate but an enabler of goal-directed, externally-focused problem solving and the formation of long-distance semantic links and structural analogies (Buckner *et al.*, 2008; Kounios and Beeman,

2009). Default network activity is typically experienced as mentalizing accompanied by social emotions. The primary hypothesis of the paper is then stated: repeated experiences of insight due to systemizing progressively and chronically decrease default-network activity, and hence both desensitize the experiencing individual to cues associated with mentalizing, including self-relevant social emotions, and enable further systemizing. Section 6 outlines predictions generated by this hypothesis that distinguish it from alternative accounts of the development of hyper-systemizing, reviews data pertaining to these predictions, and proposes a number of additional experimental tests.

2. Systemizing and hyper-systemizing in the human population

The systemizing concept was introduced by Baron-Cohen (2002) to characterize the positive capabilities commonly expressed by high-functioning individuals with ASD. The terms “systemizing” and “mentalizing” refer both to an overall orientation toward one problem-solving orientation or the other and to the activity of solving problems from that orientation. The distinction between these two problem-solving orientations drives a wedge between activities commonly thought of as explanation (Gopnik, 2000; Lombrozo, 2006), and suggests that humans possess an integrated and specialized cognitive-affective “system brain” analogous to the “social brain” employed in mentalizing (Adolphs, 1999; Dunbar, 2003; Saxe *et al.*, 2004; Frith, 2007).

Experimental instruments have been developed to measure orientation or bias toward systemizing or mentalizing as a cognitive style (Baron-Cohen *et al.*, 2003); however, domain-independent experimental instruments to quantitatively measure systemizing or mentalizing ability or capability across broad populations are not available. Hence while it is clear that some individuals – e.g. theoretical physicists – are both highly oriented toward and highly capable of systemizing while others – e.g. novelists – are both highly oriented toward and highly capable of mentalizing, the relationships between bias and ability or capability within and between systemizing and mentalizing are not well understood. Experimental studies as well as common experience indicate that a bias toward systemizing typically, but not always, predicts a bias against mentalizing and vice-versa (Baron-Cohen *et al.*, 2002; 2003). Individuals with ASD exhibit functionally-deficit mentalizing and a strong bias toward systemizing (Baron-Cohen *et al.*, 2001; 2002; 2003); but the consequences of deficit systemizing, termed “system blindness” (Baron-Cohen *et al.*, 2002), have not been investigated. Hence models that combine orientation towards and ability or capability in systemizing and mentalizing into a single cognitive dimension (e.g. Crespi and Badcock, 2008, Fig. 4) may be useful as heuristics, but both overstate and oversimplify current understanding of the positive and negative correlations between systemizing and mentalizing biases and capabilities.

Both males and females considered as subpopulations exhibit the full range of biases for or against systemizing; however, males as a group show a greater orientation toward systemizing than females as a group (Baron-Cohen *et al.*, 2003; Goldenfeld *et al.*, 2006; Nettle, 2007). Differences in testosterone exposure *in utero* have been advanced as a possible explanation of this difference (Baron-Cohen *et al.*, 2004). Consistent with a prenatal origin, the distinction between systemizing and mentalizing is observable early in development, initially in association with distinct orientations toward animate and inanimate objects (Karmaloff-Smith, 1995; Subrahmanyam *et al.*, 2002). A robust ability to attribute agency develops in infancy (Johnson *et al.*, 2007; Saxe *et al.*, 2005; 2007); infant tendencies toward mentalizing predict later childhood tendencies toward mentalizing (Wellman *et al.*, 2004). An understanding that inanimate objects, unlike animate agents, respond in predictable ways to external

causes develops in early childhood (Gopnik and Schultz, 2004). While infants separately categorize self-propelled inanimate objects and attribute internal casual powers to them (Subrahmanyam *et al.*, 2002; Luo *et al.*, 2009), an understanding that objects may exhibit spontaneous behaviors due to hidden, internal mechanical causes – arguably the earliest indication of robust systemizing – can typically be inferred around age four (Sobel *et al.*, 2007). A competitive mechanism for switching between systemizing and mentalizing strategies for problem solving is suggested by classic experiments with animated geometric figures, which demonstrate transitions between systemizing and mentalizing, in both children and adults, driven by minor changes in the motions of simple geometric shapes (Scholl and Tremoulet, 2000); such switching may be implemented by mirror-system neurons attuned to biological motion patterns (Puce & Perrett, 2003; Engel *et al.*, 2007). A bias towards mentalizing correlates strongly with the Big Five personality factor Agreeableness; a bias toward systemizing correlates significantly with the Big Five factors Conscientiousness and Openness (Nettle, 2007), consistent with documented personality characteristics of scientists, and particularly of more creative scientists (Feist, 1998).

Hyper-systemizers are individuals displaying an unusually strong to exclusive orientation toward systemizing, e. g. as indicated by a Systemizing Quotient (SQ) score more than two standard deviations above the mean (Baron-Cohen *et al.*, 2002; Baron-Cohen, 2008). Abundant anecdotal evidence suggests that high-functioning hyper-systemizers also have high-level capability in systemizing. However, this correlation has not been studied experimentally with large, unbiased populations, so “hyper-systemizer” must be regarded as indicating a strong orientation toward systemizing regardless of capability. The notion of a hyper-systemizer is culture- and even subculture-dependent: adopting a systemizing approach to human origins is unremarkable in contemporary western Europe, for example, but remarkable as a minority position in the United States (e.g. Paul, 2009); adopting a systemizing approach to human language understanding is commonplace in cognitive neuroscience or artificial intelligence circles, but is considered incoherent by many philosophers (e.g. Searle, 1980). Human beings have a well-documented tendency to over-attribute intentions and autonomous agency (Scholl & Tremoulet, 2000; Atran & Norenzayan, 2004; Rosset, 2008), a tendency that is evolutionarily adaptive in a niche populated by predators and enemies (Barrett, 2000; Boyer & Bergstrom, 2008). Historical cultures dominated by supernatural explanations for most everyday phenomena were likely far less tolerant of hyper-systemizers than is our present highly-technological culture.

Hyper-systemizing in conjunction with deficit mentalizing capability presents in early development as ASD, with severity dependent on the severity of mentalizing deficit (Baron-Cohen *et al.*, 2001; 2002; 2003; Crespi and Badcock, 2008; Ring *et al.*, 2008); physical sciences, engineering and mathematics are common career choices of high-functioning individuals diagnosed with Asperger's syndrome or high-functioning autism (Baron-Cohen *et al.*, 2001; Fitzgerald & O'Brien, 2007). Hyper-systemizing itself, however, is not generally regarded as pathological. Despite the strong correlation between hyper-systemizing and ASD (Baron-Cohen *et al.*, 2001; 2002; 2003; Goldenfeld *et al.*, 2006; Baron-Cohen, 2008), hyper-systemizing is also observed in both males and females with normal-range mentalizing capability (Baron-Cohen *et al.*, 2003; Goldenfeld *et al.*, 2006), as would be expected given the existence of highly systemizing-oriented scientists, engineers and mathematicians of both sexes who are able to attribute mental states to others, predict behavior on the basis of inferred intentions, and function effectively in large organizations. A diagnosis of ASD as a pathology, moreover, requires “clinically significant impairment” (APA, 1994). In one study of 454 university science and technology students, Baron-Cohen *et al.* (2001) report that 4.6% scored in the range suggestive of ASD on the Autism Quotient instrument, of whom 80% met objective diagnostic criteria for ASD. These

presumptive hyper-systemizers nonetheless reported no significant distress and were considered free of pathology. These observations suggest strongly that the hyper-systemizing observed in association with ASD is not a cause of ASD, but rather a developmental consequence of deficit mentalizing or other underlying deficits (Rejendran and Mitchell, 2006; Markram *et al.*, 2007). The mechanisms driving development of hyper-systemizing in the absence of pathology and in conjunction with normal-range mentalizing capability, and in particular the motivational components of such asymmetric development, are not well characterized. The next section reviews one potential contributor to the development of hyper-systemizing, the affective experience of problem-solving with insight.

3. Affective experiences associated with systemizing

Common experience as well as developmental and adult cognitive studies suggest that many if not most people enjoy solving problems; it has been suggested that the human drive to explain is universal and analogous to the human drive for sex (Gopnik, 2000). The pleasure associated with discovery and understanding is regarded as an intrinsic motivation toward learning, not only in academic environments (Gottfried, 1985), but also in curiosity-driven unstructured play (Gibson, 1988; Karmaloff-Smith, 1995; Kaplan and Oudeyer, 2007). However, common experience also suggests that there are large individual differences in the extent to which the pleasure of learning is motivational, and desires that do not have pleasurable experience as primary components are often motivational in both academic (Covington, 2000) and general (Reiss, 2004a) environments. Consistent with such diversity, problem solutions or explanations do not uniformly induce pleasure; many religious explanations, for example, appear to be designed to induce fear, dread, or a social emotion such as affiliative solidarity (Reiss, 2004b). Hence, while a drive to explain may be universal, it appears to be a drive with multiple, separable components, and with multiple affective associations.

A diverse body of evidence indicates that feelings of pleasure are both more commonly and more likely to be associated with systemizing rather than mentalizing solutions or explanations. Mentalizing capacities develop in, and are generally regarded as having evolved in, small-group social contexts in which correct assessments of the intentions of others are critical for survival (Adolphs, 1999; 2003; Dunbar, 2003). Hence solutions to the most basic mentalizing problem – does this approaching person intend help or harm? - are naturally associated with primary emotions of affiliative bonding and fear. Solutions of more subtle mentalizing problems, such as determining whether a partner in a social exchange is cheating, are typically associated with social emotions such as, in this case, righteous anger or jealousy (Adolphs, 2003). Consistent with the strong emotional associations of mentalizing, expectation-reality conflicts in mentalizing often induce anxiety, and can induce pain comparable in character and intensity to physical pain (Eisenberger and Lieberman, 2004; Eisenberger, 2006). In contrast, the inanimate objects with which most systemizing is concerned tend not – with the exception of some products of technology – to be harmful or rewarding in and of themselves in the ways that animals and other humans are. Systemizing solutions and explanations can, therefore, be expected to lack the rich emotional tonality associated with mentalizing. Consistent with this expectation, the primary emotions associated with systemizing in self reports are pleasure and frustration (Shaw, 1999; Amabile *et al.*, 2005). The dynamic range of these emotions is quite large: prominent historical figures from Archimedes onward have reported or displayed emotions ranging from mild pleasure to ecstatic elation following discoveries, and intense frustration bordering on despair when solutions seemed unreachable (Fitzgerald and O'Brien, 2007). Technical workplace subjects report feelings of pleasure ranging from “relieved and happy” to “all hyped” and “wonderful” accompanying successful

systemizing, and varying levels of frustration accompanying failures (Amabile *et al.*, 2005). Similarly, working scientists report frustration levels from “agitation” to “real bitter” in the face of seemingly intractable problems, and describe break-through insights as “really exciting” and “orgasm” (Shaw, 1999). The verbal richness of such reports indicates none of the alexithymia typically associated with ASD (Fitzgerald and Bellgrove, 2006); indeed, many subjects in these studies provide elaborate descriptions of “Aha!” contexts that clearly indicate competent mentalizing.

Personality characteristics and pathologies typical of hyper-systemizers and hyper-mentalizers (typically hypo-systemizers; Crespi and Badcock, 2008) provide supporting evidence that systemizing and mentalizing are associated with distinct affective spectra. Obsessive focus on work, social withdrawal, and neglect of personal maintenance are common to creative artists and scientists, but creative artists do not display the aversion to strong social emotions and interpersonal conflict typical of creative scientists (Feist, 1998; 1999). Highly creative scientists report bouts of intense frustration, but clinical depression is more common in highly creative writers and artists (Feist, 1999; Nettle, 2001). Highly creative scientists may experience ecstatic pleasure following insights, but also report periods of uncertainty following the initial pleasure (Shaw, 1999) and demonstrate high levels of the Big Five factor Conscientiousness (Feist, 1998, Nettle, 2007). Highly creative artists demonstrate lower levels of Conscientiousness (Feist, 1998), and creative writers and artists are more prone to mania, and to psychotic spectrum disorders (PSD) in general (Nettle, 2001). It is interesting in this regard that while many prominent scientists have been retrospectively diagnosed with ASD (Fitzgerald & O'Brien, 2007), many prominent religious figures have been retrospectively diagnosed with epilepsy (Saver and Rabin, 1997), consistent with the more general association of hyper-religiosity with PSD (Nettle, 2001; Previc, 2006; Crespi and Badcock, 2008).

While the ability of putative hyper-systemizers to experience intense pleasure following insightful discovery has been much celebrated, a second characteristic of this population is less often noted: many take and have taken substantial risks, and many have died, in pursuit of explanations. Much scientific and technical work is inherently dangerous, and the dangers are often best understood by those who risk them. Marie Curie and Enrico Fermi, for example, are only the best known of the nuclear physicists who have died of radiation-associated cancers. Risky self-experimentation by naturalists, chemists and physicians has been commonplace throughout history. The pursuit of systemizing explanations has, moreover, been vigorously suppressed by governments and religious authorities until well into the 20th century, and from Socrates to Nikolai Vavilov the list of scientists and philosophers persecuted or executed for violating legal or socio-cultural sanctions against systemizing as a cognitive activity is long. It could be argued, on a case-by-case basis, that social status, monetary benefit or other anticipated rewards motivated this history of risk taking, but the alternative hypothesis of intrinsic motivation is more consistent with cases for which self-reports are available (Csikszentmihalyi, 1996; Fitzgerald and O'Brien, 2007). An understanding of hyper-systemizing as a cognitive-affective orientation must provide a natural explanation for this extraordinary tolerance of mortal risk.

Behavior patterns characterized by intense frustration and pleasure, obsessive focus on reward-producing activities in conjunction with social withdrawal, and a high tolerance for significant risk in the pursuit of rewards are symptomatic of addictions (Hyman *et al.*, 2006). In the case of drug addictions, such symptoms reflect pathological levels of orientation toward drug-related external cues and desire for further drug experiences driven by drug-induced sensitization of the attentional and motivational systems (Robinson and Berridge, 1993; 2008). Hyper-systemizers are characterized by an attentional and motivational focus on situations and activities conducive to systemizing, and it is

reasonable to ask how this attentional and motivational focus develops. The next section reviews neurocognitive evidence that problem solving by systemizing is coupled to the mesocorticolimbic reward pathway via the task-monitoring functions of anterior cingulate cortex (ACC). Reward-pathway activation by an endogenous signal from ACC decouples the positive affect associated with systemizing from extrinsic rewards, and provides a mechanism by which a sudden “Aha!” experience can be generated by the problem-solving process itself. Evidence that progressive and chronic deactivation of the default network (Buckner et al., 2008) provides an ACC-mediated mechanism by which repeated episodes of insightful problem solving can be expected to result in attentional and motivational sensitization toward further systemizing is then reviewed in Section 5.

4. Activation of the reward pathway by systemizing

Natural rewards such as food and sex reliably induce pleasure. However, it is difficult to explain the pleasure associated with solving problems by systemizing in terms of immediate or even anticipated natural rewards. Problems typically solved by mentalizing are in most cases, and arguably in all cases relevant to early development, associated with perceptual cues signaling either natural rewards (e. g. smiling human faces) or natural dangers (e. g. angry human faces, snarling animals). In contrast, the perceptual cues that contribute to representations of problems to be solved by systemizing do not in general signal natural rewards; regularities among events or symmetries in structures, for example, are not rewarding *per se*. Anticipated extrinsic rewards associated with systemizing solutions, such as social status or economic rewards associated with successful technologies, may be distant in time, dependent on multiple external factors and hence highly uncertain, and in cultural contexts in which systemizing is frowned upon, fraught with danger. Associations between effortful activity and extrinsic reward such as those learned in school are, moreover, unlikely to be specific to systemizing. Such dissociation of systemizing from natural or other extrinsic rewards suggests that an endogenous cognitive state generated in the course of problem solving, not a perceived or imagined external reward for a successful outcome, activates the reward system and hence generates the feeling of pleasure associated with successful systemizing. That an endogenous process-specific representation, not a perceived or imagined aspect of the problem or intermediate step in its solution, triggers the pleasure response associated with insight is further suggested by the fact that the sudden, intense pleasure of “Aha!” is associated not with the often-extended process of grappling with a problem, but rather with the recognition of a solution.

Multiple lines of evidence, all indirect, suggest that ACC is the locus of interaction between problem solving processes and the mesocorticolimbic reward pathway, and that an ACC-generated task-completion signal, in an appropriate context, is the internal representation that triggers an “Aha!” response. ACC monitors conflicts between cortical representations of current and goal states and evaluates the results of actions against the goals they are intended to achieve (Botvinick *et al.*, 2004; Botvinick, 2007; Carter and van Veen, 2007), signaling both positive and negative progress toward a goal in a context-dependent way (Kennerley *et al.*, 2006; Holroyd and Coles, 2008; Quilodran *et al.*, 2008). Reciprocal interactions between ACC and dorsolateral prefrontal cortex (DLPFC) implement attentional control (Ridderinkhof *et al.*, 2004; Carter and van Veen, 2007); reciprocal interactions between ACC and orbito-frontal cortex (OFC) enable context-dependent re-evaluation of goals (Rushworth *et al.*, 2007; Rolls and Grabenhorst, 2008). ACC task-monitoring signals to the basal forebrain modulate the intensity of attentional focus; feedback of task-monitoring signals to OFC via nucleus accumbens (NAc), ventral tegmental area (VTA) enables the experience of attentional focus

combined with expectation of success as concentration, task difficulty, motivation to continue, or fatigue (Sarter *et al.*, 2006; Grace *et al.*, 2007; Boksem and Tops, 2008). Hence ACC measures progress, or lack thereof, toward a solution, and communicates this measurement to other components of the mesocorticolimbic reward pathway in a way that not only enables the use of such discrepancies to regulate attention and motivation toward problem solving, but also enables both the affective re-evaluation of goals and the affective experience of problem-solving progress or frustration.

Consistent with the role of ACC as an affect-mediating task-progress monitor, processes in which actual-state to goal-state conflicts are small are experienced as pleasurable and motivating, presumably via activation of dopamine-driven reward pathways (Kaplan and Oudeyer, 2007). Hedonic experience is generated by activation of opiate and cannabinoid receptors in NAc and ventral pallidum, but is experienced as associated with sensory or imaginative input and influences attentional focus and motivation toward goals via its representation in OFC and ACC (Kringelbach and Berridge, 2009; Smith *et al.*, 2009). Fluent performance of an effortful task requiring focused concentration can be intensely enjoyable, provided the performance remains fluent and conscious decision-making is not required (Csikszentmihalyi, 1996; Dietrich, 2004); the fact that the effortful activity itself, not just its successful outcome, is experienced as pleasurable indicates that in such cases process monitoring is directly coupled to affective experience. The specific dissociation of executive control and the feeling of effort in a patient with left ACC and OFC damage (Naccache *et al.*, 2005) confirms this direct link from ACC to the conscious representation of task progress. Positive affect correlates with ease of performance, and hence presumably with small actual-state to goal-state conflict signals from ACC, in tasks in which performance difficulty is varied without a subject's knowledge (Winkielman and Cacioppo, 2001; Winkielman *et al.*, 2003), and processing fluency may explain preferences for well-known prototypes in cases ranging from facial features to abstract designs (Winkielman *et al.*, 2006). These observations all suggest that positive affect increases as ACC-signaled discrepancy from a problem-solving goal decreases. If this is correct, the sudden convergence of a problem-solving process to its goal would be expected to induce a strong task-completion signal from ACC, and via its connections to NAc and VTA a burst of positive affect, an “Aha!” experience.

Activation of the mesocorticolimbic reward pathway by drugs of abuse confers enhanced salience on drug-related cues and sensitizes the attentional and motivational systems toward further drug use, leading in susceptible cases to drug addictions (Robinson and Berridge, 1993; Kelley and Berridge, 2002; Hyman, 2005; Nestler, 2005; Hyman *et al.*, 2006; Grace *et al.*, 2007; Robinson and Berridge, 2008). While no data comparing levels of reward-pathway activation by drugs and by endogenous task-completion signals from ACC are available, both reported intensities of positive affect accompanying insights and the often obsessive and risk-tolerant attentional and motivational orientation toward systemizing displayed by hyper-systemizers suggest that reward-driven sensitization mechanisms may also be activated by strong ACC task-completion signals. Such mechanisms can be expected to confer enhanced salience on solution-related external or endogenous cues and to sensitize the attentional and motivational systems to detect and respond to such cues. Attentional and motivational sensitization toward external cues such as quiet or the familiar confines of a personal workspace may contribute to some of the common personality characteristics of hyper-systemizers. The primary hypothesis of the present paper is, however, that repeated episodes of insightful problem solving result in attentional and motivational sensitization toward further systemizing via a particular mechanism: progressive and chronic deactivation of the default network (Buckner *et al.*, 2008). The next section reviews evidence that ACC-mediated default-network deactivation is not only a consistent correlate but an enabler of insightful solutions of systemizing problems. It shows that the hypothesis of

progressive and chronic default-network deactivation not only provides an explanation for the typical bias against mentalizing displayed by hyper-systemizers without pathology; it also provides a mechanism by which hyper-systemizers become better analogical reasoners and hence better problem solvers.

5. Default-network deactivation as an enabler of systemizing.

Three aspects of context and interoceptively-accessible cognitive “set” are consistent across anecdotal reports, survey-based measurements, and experimental studies of systemizing resulting in insight: moderate social withdrawal (Feist, 1998; Fitzgerald and O'Brien, 2007), positive affect (Shaw, 1999; Fredrickson, 2004; Amabile et al., 2005, Fitzgerald and O'Brien, 2007; Kounios and Beeman, 2009), and an attentional focus correlated with decreased default-network activation (Buckner *et al.*, 2008; Kounios and Beeman, 2009). These effects are causally related. The default network links temporal-parietal junction (TPJ) areas implementing ToM to attention-control areas of medial prefrontal cortex (PFC) including ACC (Raichle and Snyder, 2007; Buckner *et al.*, 2008). Default network activity is experienced as self-conscious and self-relevant reminiscence and future-oriented planning, typically in the modality of inner speech and accompanied by content-relevant emotions, mainly social emotions (Northoff *et al.*, 2006; Raichle and Snyder, 2007; Buckner *et al.*, 2008; Schilbach *et al.*, 2008). Default network activity is often obsessive, and high levels of default activity correlate with both major depression (Sheline et al., 2009) and schizophrenia (Kim et al., 2009). Artists often find inspiration in such emotionally-colored reflections (Nettle, 2001), but scientists tend toward discomfort with social emotions, particularly those associated with inter-personal conflict (Feist, 1998), and employ social withdrawal both to escape such emotions and to avoid distractions. Scientists are often happiest when “lost in their work,” a flow-like state characterized by attentional focus on non-self-oriented elements of a task, and hence correlated with low default network activation (Buckner *et al.*, 2008). It seems reasonable, therefore, to propose that decreased default network activity, experienced as a pleasant absence of uncomfortable and distracting self-oriented social emotions, is associated with insight by hyper-systemizers, and is rendered both attentionally salient and motivationally desirable by sensitization.

Decreased default-network activity, and hence decreased mentalizing, self-referential thinking, and experienced social emotions (Buckner *et al.*, 2008) appears to be not only a correlate of problem solving by insight, but an enabler and possibly a prerequisite of problem solving by insight. Experimental studies of insight have focused on the formation of distant semantic connections (Bowdon *et al.*, 2005), a precursor of the formation of the structural analogies typical of insightful solutions of real-world problems (Gentner, 2003; 2005; Leech *et al.*, 2008). Formation of distant semantic connections requires activity in predominantly right-hemisphere association areas (Jung-Beeman *et al.*, 2004; Kounios *et al.*, 2007; Sandkuhler and Bhattacharya, 2008) that overlap strongly with the default network (Bar, 2008). Recruiting these resources for systemizing would, therefore, require suppressing attention to default-network activity. Interactions between ACC and rostral PFC appear to implement this suppression of default-network activity in insight (Kounios *et al.*, 2006; Subramaniam *et al.*, 2009), analogy (Green *et al.*, 2006), and integration of solution components from different subtasks (De Pisapia and Braver, 2008). ACC-mediated suppression of default-network activity correlates with and may implement the facilitation of insight by positive affect (Subramaniam *et al.*, 2009; Kounios and Beeman, 2009). The involvement of rostral PFC in suppression of default-network activity is consistent with its general role in attention switching and multi-tasking (Gilbert *et*

al., 2005; Dreher *et al.*, 2008).

Decreased default-network activity appears to be a general correlate of hyper-systemizing. High-functioning individuals with ASD tend to be hyper-systemizers (Baron-Cohen, 2002; Baron-Cohen *et al.*, 2002; 2003); default-network activity is significantly lowered in ASD, as is the relative decrease in activity accompanying default-network deactivation in attention-switching tasks (Kennedy *et al.*, 2006). Individuals with ASD exhibit low levels of self-referential thinking (Lombardo *et al.*, 2007) and functional differences in medial-frontal connection patterns that overlap the default network (Gilbert *et al.*, 2009), both suggestive of reduced default-network activity as a correlate of (in this case pathology-associated) hyper-systemizing. Hyper-systemizers without pathology appear more capable of maintaining a high degree of attentional focus on non-self-relevant stimuli (Billington *et al.*, 2008), consistent with the obsessive attention to patterns typical of ASD (Baron-Cohen, 2002) and the obsessive attention to their work typical of creative scientists (Feist, 1998). Maintenance of externally-focused, non-self-relevant attention requires default-network deactivation (Buckner *et al.*, 2008). Highly-capable scientists, who tend as a group to be more oriented toward systemizing, are both more likely to engage in and more likely to be successful using analogical reasoning (Feist & Gorman, 1998), an activity enabled by decreased default-network activity (Green *et al.*, 2006). While these observations do not establish causation, they are consistent with decreased default-network activity and hence reduced mentalizing being a pre-requisite for successful systemizing.

The foregoing considerations all suggest that if the mesocorticolimbic reward pathway is activated by task-completion signals generated by ACC in response to sudden, insightful solutions to systemizing problems, the activity of this pathway can be expected to 1) increase any pre-existing bias toward systemizing; 2) decrease the frequency and intensity of mentalizing, self-referential thinking, and self-relevant emotional experience; and 3) increase the probability of recognizing distant semantic connections and analogies, all by the common mechanism of ACC-mediated suppression of default-network activity. Reciprocal ACC – OFC connections implement the cognitive representation of hedonic experience, and mediate its effects on attention and motivation (Kringelbach and Berridge, 2009; Smith *et al.*, 2009). While the locus or loci of reward-driven long-term sensitization of attentional and motivational responses in drug addictions (Robinson & Berridge, 1993; 2008) has not been identified, encodings of reward-related experience by the ACC – OFC loop presumably play at least a facilitating role in such sensitization. As ACC is known to suppress default-network activity in insight (Kounios *et al.*, 2006; Subramaniam *et al.*, 2009), analogy (Green *et al.*, 2006), and solution integration (De Pisapia and Braver, 2008), it is reasonable to suppose that ACC may couple solution-related hedonic representations to solution-related default-network deactivation in any or all of these cases. The hypothesis that hyper-systemizing is a result of ACC-mediated default-network deactivation in response to insight provides a mechanism by which hyper-systemizing can develop in individuals with normal-range mentalizing capability, and explains both an increasing bias against mentalizing and increasing analogical and integrative ability and hence increased capability in solving systemizing problems in such individuals. Moreover, it explains common asocial personality characteristics of hyper-systemizers, including their obsessive attention to their work, tendencies toward social withdrawal and aversion to social conflict and social distractions. It is significant, moreover, that this ACC-mediated mechanism would not be expected to develop hyper-mentalizing from a pre-existing mentalizing bias. As discussed in Section 3, mentalizing is associated with a broad spectrum of social emotions, many of them unpleasant; hence mentalizing cannot be expected to consistently activate the reward pathway. Predictions that distinguish the default-network deactivation model of hyper-systemizing from alternative models, available data pertaining to them, and experiments that would

further test them are discussed in the next section.

6. Predictions of the default-network deactivation model of hyper-systemizing

The model of hyper-systemizing as a result of ACC-mediated default-network deactivation in response to insight developed above makes three general predictions. First, it predicts that hyper-systemizing is neither congenital nor exclusively an outcome of deficit mentalizing, as is observed in ASD, but rather that hyper-systemizing can develop in susceptible individuals exhibiting normal-range mentalizing capability. Second, it predicts that both the developmental progression and adult expression of hyper-systemizing are primarily dependent on endogenous rewards, and particularly on the endogenous experience of insightful solutions to systemizing problems. While the model does not predict that learned associations with natural or other extrinsic rewards are irrelevant, particularly in early development, it does predict that extrinsic rewards play a progressively smaller role as an initial orientation toward systemizing develops into hyper-systemizing. Third, the model predicts that the development of hyper-systemizing is driven by a specific causal feedback loop: attentional and motivational orientation toward cues indicative of systemizing problems together with systemizing ability increase the probability of experiencing insightful solutions to systemizing problems, hedonic responses to insight cause default-network deactivation, and default-network deactivation desensitizes individuals to mentalizing-indicating cues, sensitizes them to systemizing-indicating cues, and further enables insightful solutions to systemizing problems. Thus the model predicts that hyper-systemizing develops in neurocognitively normal individuals with an initial moderate orientation toward and ability in systemizing who both experience insightful problem solutions and experience such solutions as pleasurable. Hyper-systemizing, on this model, is not a pathology, but a natural developmental outcome from a particular set of initial circumstances.

The prediction that hyper-systemizing can develop in the presence of normal-range mentalizing distinguishes the default-network deactivation model from congenital-cause models, such as that of Crespi and Badcock (2008), that characterize hyper-systemizing as an obligate developmental correlate or consequence of deficit mentalizing. The default-network deactivation model does not attempt to explain the origin of an initial bias toward systemizing in some individuals; such a bias may be innate, or may develop in response to differential experience with objects during infancy (Rakison and Yermolayeva, 2010). What the default-network deactivation model provides is a mechanism for amplifying a small initial bias toward systemizing into the strong bias characteristic of hyper-systemizing. The operation of this mechanism is independent of mentalizing capability; hence the default-network deactivation model predicts both that hyper-systemizers would display a distribution of mentalizing capabilities extending well into the normal range, and that the development of hyper-systemizing would not significantly negatively impact mentalizing capability, even if it significantly reduced attentional and motivational responses to mentalizing cues and hence expression of mentalizing. The identification of hyper-systemizers within populations of “normal” subjects (Baron-Cohen *et al.*, 2002; 2003; Nettle, 2007) confirms the first of these predictions. Joint measurements of SQ and mentalizing capability, as opposed to mentalizing bias, in “normal” subjects at developmental timepoints from childhood to adulthood would test the second.

The prediction that hyper-systemizing develops by an endogenously-driven sensitization mechanism distinguishes the default-network deactivation model from conventional learning models, still dominant in most formal educational and management settings, that emphasize learned associations with natural

or other extrinsic rewards over intrinsic motivation (e.g. Covington, 2000). Association of systemizing as an activity with actual or anticipated extrinsic rewards involves a form of self-relevant thinking, and hence would be expected to activate the default network (Buckner *et al.*, 2008) and decrease both the ability to maintain external focus (Buckner *et al.*, 2008; Kounios & Beeman, 2009) and the likelihood of successful analogies (Green *et al.*, 2006), solution integrations (De Pisapia and Braver, 2008), and insights (Kounios *et al.*, 2006; Subramaniam *et al.*, 2009). Awareness or contemplation of extrinsic rewards would, therefore, be expected to decrease, not increase, systemizing performance in real time. Empirical and meta-analytic studies of the short-term effects of extrinsic motivation on academic and workplace tasks have not been conclusive; Deci *et al.* (1999) find evidence of a general negative impact of extrinsic rewards on performance based on a meta-analysis of three decades of laboratory studies, while Cameron *et al.* (2001) using a similar meta-analysis find evidence for negative impacts only for particular reward procedures. Existing studies have not, however, controlled for subject SQ, and have not focused on tasks designed to require default-network deactivation. The default-network deactivation model would predict that expected extrinsic rewards would become increasingly ineffective or counterproductive with increasing subject SQ, and that this effect would be more pronounced on tasks designed to require long-distance semantic connections, analogy, or partial-solution integration. It would also predict that extrinsic rewards presented as distractors would disrupt performance on systemizing tasks more than on mentalizing tasks, and would disrupt the performance of high-SQ more than low-SQ individuals.

A second strand of evidence relevant to the question of intrinsic versus extrinsic rewards is provided by the intense and obsessive interests sometimes developed by preschool children. Intense childhood interests often involve functional objects with discernible parts, i.e. involve systemizing, and occur most frequently in males. Children who develop such interests typically do so suddenly and without unusually-high extrinsic rewards from parents or others as a response; indeed children with intense preschool interests often abandon them due to negative reinforcement from teachers or peers (DeLoach *et al.*, 2007). Prospective longitudinal studies that assessed the late-childhood or early-adulthood SQ and career choices of children exhibiting intense systemizing interests, as well as retrospective studies of childhood interests of individuals selected only on the basis of SQ score would be expected on the default-network deactivation model to yield positive correlations between childhood and adult systemizing and negative correlations with mentalizing in either temporal direction. These correlations would, moreover, be expected to be independent of extrinsic rewards either supporting or discouraging systemizing within the ranges typically experienced by children or young adults.

The prediction that the development of hyper-systemizing is driven by feedback between experiences of insight and default-network deactivation distinguishes the default-network deactivation model from alternative models that rely primarily on intrinsic motivation, including in particular the idea, widespread implicitly if not explicitly in both the academic and popular literature, that interest-driven selection of subject-specific input to a general-purpose learning mechanism is sufficient to generate high-level subject-specific orientation and expertise. This prediction has three components, one correlational and two causal. First, the model predicts that hyper-systemizing in the absence of pathology (e.g. ASD) is positively correlated with systemizing capability, but uncorrelated with mentalizing capability. The mechanisms proposed by the model are, moreover, systemizing-specific; the model is consistent with a lack of correlation between hyper-mentalizing and mentalizing capability. Second, the model predicts that chronically decreased default-network activity causes desensitization to cues indicating mentalizing, and causes sensitization to cues indicating systemizing. Third, the model predicts that chronically decreased default-network activity causes increased

systemizing capability by enabling distant semantic connections, analogical reasoning and solution integration. This effect is predicted to be systemizing-specific; the model is consistent with an absence of facilitation of these problem-solving capabilities in mentalizers.

That high levels of systemizing capability positively correlates with hyper-systemizing seems obvious: one need only examine typical scientists, technologists, engineers and mathematicians. Studies of this correlation in samples identified only by high SQ scores, however, remain rare. The recent observation that high systemizing bias correlates with high mental rotation ability (Cook and Saucier, 2010) is consistent with the prediction. However, it remains possible that a substantial number of pathology-free hyper-systemizers who are nevertheless poor at solving systemizing problems stated at their own level of knowledge and understanding may exist. Identification of such individuals would substantially challenge the current model, as it would require that such individuals constituted “special cases” that either did not experience insights, or did not experience insight as pleasurable. The model prediction that hyper-systemizing correlates with enhanced systemizing ability can be tested in an unbiased way by screening subjects in problem-solving ability studies that probe systemizing skills for systemizing versus mentalizing bias. Hyper-systemizers would be expected to perform better than “balanced” or mentalizer controls of matched verbal ability on semantic-association tests as employed in insight studies (Jung-Beeman *et al.*, 2004; Kounios *et al.*, 2007; Sandkuhler and Bhattacharya, 2008) and on standard verbal analogies (Gentner, 2003; Holyoak, 2005). Hyper-systemizers would also be expected to perform better than balanced or mentalizer controls of matched visuo-spatial ability on tool-improvisation problems, which are effectively visuo-spatial analogies (Fields, 2010). Conversely, hyper-mentalizers would be expected to perform no better than balanced or systemizer controls on mentalizing tasks, such as correctly inferring the intentions of actors in stories.

The prediction that decreased default-network activity causes both desensitization to mentalizing cues and sensitization to systemizing cues can be tested both mechanistically and at the level of outcomes. A straightforward consequence of this prediction is that individuals strongly oriented to both systemizing and mentalizing cues should be rare, as indeed they are observed to be (Goldenfeld *et al.*, 2006; Crespi & Badcock, 2008). Individuals with chronically-decreased default-network activity due to pathology (i.e. ASD) should be highly oriented toward systemizing and insensitive to mentalizing cues, as indeed they are (Gilbert *et al.*, 2009). A further consequence is that an orientation toward systemizing, once developed, should persist in the same way that pathological “wanting” for drugs persists in addicts (Robinson & Berridge, 1993; 2008). Hyper-systemizers should be unlikely to abandon systemizing even after prolonged periods without significant insights. This prediction appears plausible in the case of scientists, but needs an unbiased test on a sample identified only by high SQ.

Priming designs in which default-network deactivation is induced by verbal or other input and confirmed by functional imaging or other means, and the effects of such deactivation on orientation toward cues are then tested, may be the most practical approach to testing the prediction of opposing sensitization and desensitization at the causal level. It is well-known that priming with stories or images that induce self-relevant emotions such as fear increase orientation towards mentalizing: telling children ghost stories will induce them to interpret every night-time sound as a monster. Tasks requiring sustained external focus reliably deactivate the default network (Buckner *et al.*, 2008), but the effectiveness of such procedures as primes is unknown. Designs employing transcranial magnetic stimulation (TMS) to suppress default-network activity may be a practical alternative to priming, provided the suppression is sufficiently specific and sustained.

The prediction that decreased default-network activity causally and specifically enables problem solving by systemizing is supported by a variety of correlational data in addition to the studies of insight (Kounios *et al.*, 2006; Subramaniam *et al.*, 2009), analogy (Green *et al.*, 2006) and solution integration (De Pisapia and Braver, 2008) already discussed. Functional imaging and TMS methods have been used extensively to characterize the large-scale networks that implement the “social brain” and hence mentalizing as a problem solving strategy (Saxe *et al.*, 2004; Frith, 2007; Crespi & Badcock, 2008). The outcome-level negative correlation between systemizing and mentalizing render these studies relevant to the characterization of systemizing. Anatomical areas associated with the default network, particularly components of TPJ, are known to overlap areas implicated in the ventral attention network (Decety and Lamm, 2007; Corbetta *et al.*, 2008; Mitchell, 2008), as well as with association areas involved in planning, semantic association, and analogy (Bar, 2008). Mirror-neuron system components located in or near TPJ are activated during mentalizing (Saxe *et al.*, 2004; Frith, 2007; Crespi & Badcock, 2008), but are also known to be responsive to non-biological as well as biological motion patterns (Vingerhoets *et al.*, 2002; Schubotz and van Cramon, 2004; Engel *et al.*, 2007), to change their specificities in response to usage patterns (Catmur *et al.*, 2007; 2008), and to be involved in such typical systemizing tasks as estimating forces in physical interactions (Hegarty, 2004; Wolff, 2007) and solving mathematical problems (Qin *et al.*, 2004; Cantlon *et al.*, 2006). These observations suggest that systemizing and mentalizing compete, over developmental time if not in real time, for temporal and parietal pattern-recognition and association resources as well as for memory, attention and motivation. To the extent that default-network activity involves a use of such resources for mentalizing, its suppression over developmental time or in real time would make these resources available for systemizing. Daydreaming provides a passive alternative to problem solving for investigating such competition for resources. Daydreaming typically involves self-relevant thoughts and involves co-activation of the default network, ACC, and rostral PFC (Christoff *et al.*, 2009); however, daydreaming studies to date have not separated subjects by systemizing or mentalizing bias. Competition for resources would be indicated if overlapping patterns of activation but significantly less self-relevant content characterized the daydreams of hyper-systemizers compared to mentalizers.

Direct experimental tests of causality from default-network deactivation to the enabling of problem solving by systemizing would require development of a reliable procedure for sustained and specific suppression of default-network activity as discussed above. To be applicable for the study of problem solving as opposed to orientation, it would have to be demonstrable, on a subject-by-subject basis, that network components shared by mentalizing and systemizing were not inhibited. A more thorough understanding of individual differences in usage of temporal, parietal, and frontal resources during problem solving may be required to develop specific inhibitors of default-network activity that are adequate to this task.

7. Conclusion

A minority of the human population exhibits a strong bias toward systemizing, a problem-solving and explanatory orientation that relies on hypothesized physical mechanisms or formal procedures, not on attributions of autonomous intentional agency. Most scientists, technologists, engineers and mathematicians, and hence most readers of this paper, are members of this minority. Some members of this “hyper-systemizing” minority suffer clinically-significant symptoms and are diagnosed with ASD; most do not. Neither the neurofunctional differences that distinguish hyper-systemizers from other people, nor the developmental processes that result in dominance of the “system brain” over the “social

brain” are well understood.

This paper offers an integrative account of the development of hyper-systemizing based on the observation that insight is both a motivator and an individually and socially significant consequence of the pursuit of systemizing explanations. The primary hypothesis of this account is that repeated episodes of insightful problem solving result in attentional and motivational sensitization toward further systemizing via progressive and chronic deactivation of the default network. This hypothesis is distinct from and inconsistent with competing ideas that hyper-systemizing is an innate orientation, or that it develops as a result of general-purpose learning processes motivated by either extrinsic or intrinsic rewards. The present hypothesis is supported by a variety of data from clinical, observational, experimental, and neurofunctional studies. It provides a mechanistic explanation for both the asociality and risk tolerance characteristic of hyper-systemizers and the apparent positive correlation between strong orientation toward and high ability in systemizing. It generates a number of empirical predictions, some of which are amenable to existing experimental or observational procedures while others require the development of new and highly specific methods for the laboratory manipulation of default-network activity.

The model of hyper-systemizing proposed here implies that hyper-systemizing is an instance of developmental specialization producing functionally-significant neurocognitive diversity across the human population. Spikins (2009; 2010) has suggested, based in part on the culturally-significant abilities associated with Asperger's Syndrome, that neurocognitive diversity may be a major driver of human evolution. For this to be the case, culturally-significant neurocognitive diversity must be both heritable and culturally tolerated. If the model proposed here is correct, hyper-systemizing will appear in human populations provided small, early-developing biases toward systemizing occur in the population, and these biases are amplifiable by the mesocorticolimbic reward pathway. In this case, the requirement for heritability is reduced to a requirement for a population-level distribution in early systemizing bias, and the requirement for tolerance is reduced to a requirement for tolerance of small biases toward systemizing. Hyper-systemizing could appear, in such a population, without assortative mating between individuals biased toward systemizing, and even without cultural tolerance for hyper-systemizing. Similar mechanisms capable of amplifying small differences in network activity into culturally-significant neurocognitive diversity may be active in other areas as well.

Acknowledgements

Thanks to Spencer Dew and Eric Dietrich for discussions that contributed to the ideas developed in this paper, and to Penny Spikins for a preprint of Spikins (2010). Three anonymous referees provided suggestions that significantly clarified the presentation.

Conflict of Interest Statement

The author declares that he has no conflicts of interest relevant to the work presented in this paper.

References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences* 3(12), 469-479.
- Adolphs, R. (2003). Cognitive neuroscience of human social behavior. *Nature Reviews Neuroscience* 4, 165-178.
- Amabile, T. M., Barsade, S. G., Mueller, J. S. & Staw, B. M. (2005). Affect and creativity at work. *Administrative Science Quarterly* 50, 367-403.
- American Psychological Association (APA) (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)*. Washington, DC: APA.
- Atran, S. & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences* 27, 713-730.
- Bar, M. (2008). The proactive brain: Using analogies and associations to derive predictions. *Trends in Cognitive Sciences* 11(7), 280-289.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences* 2(2), 248-254.
- Baron-Cohen, S. (2008). Autism, hypersystemizing, and truth. *The Quarterly Journal of Experimental Psychology* 61(1), 64-75.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders* 31, 5-17.
- Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R. & Hill, J. (2002). The exact mind: Empathizing and systemizing in autism spectrum conditions. In U. Goswami (Ed.) *Handbook of Cognitive Development*. Oxford: Blackwell.
(http://www.autismresearchcentre.com/docs/papers/2002_BCetal_Goswami.pdf)
- Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N. & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London B* 358, 361-374.
- Baron-Cohen, S., Knickmeyer, R. C. & Belmonte, M. K. (2004) Sex differences in the brain: Implications for explaining autism. *Science* 310, 819-823.
- Barrett, J. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences* 4, 29-34.
- Berridge, K. E. & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* 199, 457-480.
- Billington, J., Baron-Cohen, S. & Bor, D. (2008). Systemizing influences attentional processes during

- the Navon task: An fMRI study. *Neuropsychologia* 46, 511-520.
- Boksem, M. A. S. & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews* 59, 125-139.
- Botvinick, M., Cohen, J. & Carter, C. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences* 8(12), 539-546.
- Botvinick, M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience* 7(4), 356-366.
- Bowden, E. M., Jung-Beeman, M., Fleck, J. & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences* 9(7), 322-328.
- Boyer, P. & Bergstrom, B. (2008). Evolutionary perspectives on religion. *Annual Review of Anthropology* 37, 111-130.
- Buckner R, Andrews-Hanna J & Schacter, D (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences* 1124, 1-38.
- Burgess, P. W., Simons, J., Dumontheil, I. & Gilbert, S. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. In J. Duncan, L. Phillips & P. McLeod (Eds.) *Measuring the Mind: Speed, Control, and Age* (pp. 217-248). Oxford University Press.
- Cameron, J., Banko, K. M. & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *Behavior Analyst* 24, 1-44.
- Cantlon, J., Brannon, E., Carter, E. & Pelphrey, K. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLOS Biology* 4(5) 0844-0854.
- Carter, C. S. & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience* 7(4), 367-379.
- Catmur, C., Walsh, V. & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology* 17, 1527-1531.
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M. & Heyes, C. (2008). Through the looking glass: Counter-mirror activation following incompatible sensorimotor learning. *European Journal of Neuroscience* 28, 1208-1215.
- Christoff, K., Gordon, A. M., Smallwood, J., Smith, R. & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences USA* 106(21), 8719-8724.
- Cook, C. M. & Saucier, D. M. (2010). Mental rotation, targeting ability and Baron-Cohen's empathizing-systemizing theory of sex differences. *Personality and Individual Differences* (in press).

- Corbetta, M., Patel, G. & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron* 58(3), 306-324.
- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology* 51, 171-200.
- Crespi, B. & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences* 31, 241-320.
- Csikszentmihalyi, M. (1996). *Creativity*. New York: Harper Perennial.
- De Pisapia, N. & Braver, T. S. (2008). Preparation for integration: The role of anterior prefrontal cortex in working memory. *NeuroReport* 19(1), 15-19.
- Decety, J. & Lamm, C. (2007). The role of right temporoparietal junction in social cognition: How low-level computational processes contribute to meta-cognition. *The Neuroscientist* 13(6), 580-593.
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125(6), 627-668.
- DeLoach, J. S., Simcock, G. & Macari, S. (2007). Planes, trains, automobiles – and tea sets: extremely intense interests in very young children. *Developmental Psychology* 43(6), 1579-1586.
- den Ouden, H. E. M., Frith, U., Frith, C. & Blakemore, S.-J. (2005). Thinking about intentions. *NeuroImage* 28, 787-796.
- Dichter, G. & Belger, A. (2007). Social stimuli interfere with cognitive control in autism. *Neuroimage* 35(3) 1219-1230.
- Dietrich, A. (2004). Neurocognitive mechanisms underlying the experience of flow. *Consciousness and Cognition* 13, 746-761.
- Dreher, J.-C., Koechlin, E., Tierney, M. & Grafman, J. (2008). Damage to the fronto-polar cortex is associated with impaired multitasking. *PLOS One* 3(9), e3227.
- Dunbar, R. (2003). The social brain: Mind, language, and society in evolutionary perspective. *Annual Reviews of Anthropology* 32, 163-181.
- Eisenberger, N. I. & Lieberman, N. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in Cognitive Sciences* 8(7), 294-300.
- Eisenberger, N. I. (2006). Identifying the neural correlates underlying social pain: Implications for developmental processes. *Human Development* 49, 273-293.
- Engel, A., Burke, M., Fiehler, K., Bien, S. & Rosler, F. (2007). How moving objects become animated: The human mirror system assimilates non-biological movement patterns. *Social Neuroscience* 3, 368-387.

- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review* 2(4), 290-309.
- Feist, G. J. (1999). Affect in artistic and scientific creativity. In S. W. Russ (Ed) *Affect, Creative Experience and Psychological Adjustment* (pp. 93-109). London: Taylor & Francis.
- Feist, G. J. & Gorman, M. E. (1998). The psychology of science: Review and integration of a nascent discipline. *Review of General Psychology* 2(1), 3-47.
- Fields, C. (2010). Implementation of structure mapping by event-file binding and action planning: A model of tool improvisation analogies. *Psychological Research* (in press) DOI=10.1007/s00426-010-0290-7.
- Fitzgerald, M. & Bellgrove, M. A. (2006). The overlap between alexithymia and Asperger's Syndrome. *Journal of Autism and Developmental Disorders* 36(4) 573-576.
- Fitzgerald, M. & O'Brien, B. (2007). *Genius Genes: How Asperger Talents Changed the World*. Shawnee Mission, KS: Autism Asperger Publishing.
- Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society of London B* 359, 1367-1377.
- Frith, C. (2007). The social brain? *Philosophical Transactions of the Royal Society of London B* 362, 671-678.
- Frith, U. and Frith, C. (1999). Interacting minds – A biological perspective. *Science* 286, 1692-1695.
- Frith, U. & Frith, C. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B* 358, 459-473.
- Gentner, D. (2003). Why we're so smart. In: D. Gentner & S. Goldin-Meadow (Eds.) *Language and Mind: Advances in the Study of Language and Thought*. Cambridge, MA: MIT Press. 195-235.
- Gentner, D. (2005). The development of relational category knowledge. In: L. Gershkoff-Stowe & D. Rakison (Eds) *Building object categories in developmental time* (pp. 245-275). Hillsdale, NJ: Erlbaum.
- Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology* 39, 3-41.
- Gilbert, S., Frith, C. & Burgess, P. (2005). Involvement of rostral prefrontal cortex in selection between stimulus-oriented and stimulus-independent thought. *European Journal of Neuroscience* 21, 1423-1431.
- Gilbert, S. J., Meuwese, J. D. I., Towgood, K. J., Frith, C. D. & Burgess, P. W. (2009). Abnormal functional specialization within medial prefrontal cortex in high-functioning autism: a multi-voxel

similarity analysis. *Brain* 132, 869-878.

Goldenfeld, N., Baron-Cohen, S. & Wheelwright, S. (2006). Empathizing and systemizing in males, females, and autism. *Clinical Neuropsychiatry* 2, 338-345.

Gopnik, A. (2000) Explanation as orgasm and the drive for understanding: The evolution, function and phenomenology of the theory-formation system. In F. Kiel & R. Wilson (Eds.) *Cognition and Explanation* (pp. 299-323). Cambridge, MA: MIT Press.

Gopnik, A. & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences* 8(8) 371-377.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology* 77(6), 631-645.

Grace, A. A., Floresco, S. B., Goto, Y. & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behavior. *Trends in Neurosciences* 30(5), 220-227.

Green, A., Fugelsang, J., Kraemer, D., Shamosh, N. & Dunbar, K. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research* 1096, 125-137.

Grimm, S., Ernst, J., Boesiger, P., Schuepbach, D., Hell, D., Boeker, H. & Northoff, G. (2009). Increased self-focus in major depressive disorder is related to neural abnormalities in subcortical-cortical midline structures. *Human Brain Mapping* 30(8), 2617-2627.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences* 8 (6) 280-285.

Holroyd, C. B. & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex* 44, 548-559.

Holyoak, K. (2005). Analogy. In: K. Holyoak and R. Morrison (Eds) *The Cambridge Handbook of Thinking and Reasoning*. (pp. 117-142) Cambridge: Cambridge University Press.

Hyman, S. E. (2005). Addiction: A disease of learning and memory. *American Journal of Psychiatry* 162, 1414-1422.

Hyman, S. E., Malenka, R. C. & Nestler, E. J. (2006). Neural mechanisms of addiction: The role of reward-related learning and memory. *Annual Review of Neuroscience* 29, 565-598.

Johnson, S. C., Shimizu, Y. A. & Ok, S.-J. (2007). Actors and Actions: The role of agent behavior in infants' attributions of goals. *Cognitive Development* 22(3), 310-322.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arumbel-Liu, S., Greenblatt, R., Reber, P. J. & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLOS Biology* 2(4), 0500-0510.

- Kaplan, F. & Oudeyer, P.-Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience* 1(1), 225-236.
- Karmaloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Kelley, A. E. & Berridge, K. C. (2002). The neuroscience of natural rewards: Relevance to addictive drugs. *Journal of Neuroscience* 22(9), 3306-3311.
- Kennedy, D., Redcay, E. & Courchesne, E. (2006). Failing to deactivate: Resting functional abnormalities in autism. *Proceedings of the National Academy of Sciences USA* 103(21) 8275-8280.
- Kennerley, S. W., Walton M. E., Behrens, T. E. J., Buckley M. J. & Rushworth M. F. S. (2006) . Optimal decision making and the anterior cingulate cortex. *Nature Neurosciences* 9(7), 940-947.
- Kim, D. I., Manoach, D. S., Mathalon, D. H., Turner, J. A., Mannell, M., Brown, G. G., Ford, J. M., Gollub, R. L., White, T., Wible, C., Belger, A., Bockholt, H. J., Clark, V. P., Lauriello, J., O'Leary, D., Mueller, B. A., Lim, K. O., Andreasen, N., Potkin, S. G. & Calhoun, V. D. (2009). Dysregulation of working memory and default-mode networks in schizophrenia using independent component analysis, an fBIRN and MCIC study. *Human Brain Mapping* 30(11), 3795-3811.
- Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B. & Jung-Beeman, M. (2006). The prepared mind: Neural activity prior to problem presentation predicts subsequent solution by sudden insight. *Psychological Science* 17, 882-890.
- Kounios, J., Fleck, J. I., Green, D. L., Payne, L., Stevenson, J. L., Bowden, E. M. & Jung-Beeman, M. (2007). The origins of insight in resting-state brain activity. *Neuropsychologia* 46, 281-291.
- Kounios, J. & Beeman, M. (2009). The “Aha!” moment: The cognitive neuroscience of insight. *Current Directions in Psychological Science* 18(4), 210-216.
- Kringelbach, M. L. & Berridge, K. E. (2009). Toward a functional neuroanatomy of pleasure and happiness. *Trends in Cognitive Sciences* 13(11), 479-487.
- Leech, R., Mareshal, D. & Cooper, R. (2008) Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences* 31, 357-378.
- Lombardo, M., Barnes, J., Wheelwright, S. & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLOS One* 9 (e883) 1-11.
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences* 10(10), 464-470.
- Luo, Y., Kaufman, L. & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology* 58, 441-486.

- Markram, H., Rinaldi, T. & Markram, K. (2007). The Intense World Syndrome – an alternative hypothesis for autism. *Frontiers in Neuroscience 1*, 77-96.
- Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex 18*, 262-271.
- Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D. & Willer, J.-C. (2005). Effortless control: executive control and conscious feeling of mental effort are dissociable. *Neurophychologia 43*(9), 1318-1328.
- Nestler, E. J. (2005). Is there a common molecular pathway for addiction? *Nature Neuroscience 8*(11), 1445- 1449.
- Nettle, D. (2001). *Strong Imagination: Madness, Creativity and Human Nature*. New York: Oxford University Press.
- Nettle, D. (2007). Empathizing and systemizing: What are they, and what do they contribute to our understanding of psychological sex differences? *British Journal of Psychology 98*, 237-255.
- Northoff, G., Heinzl, A., de Greck, R., Bernpohl, F., Dobrowolny, H. & Panksepp, J. (2006). Self-referential processing in our brain – A meta-analysis of imaging studies on the self. *NeuroImage 31*, 440-457.
- Paul G. (2009). The chronic dependence of popular religiosity upon dysfunctional social conditions. *Evolutionary Psychology 7*(3), 398-441.
- Previc, F. H. (2006). The role of extrapersonal brain systems in religious activity. *Consciousness and Cognition 15*(3), 500-539.
- Puce, A. & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Proceedings of the Royal Society of London B 358*, 435-445.
- Qin, Y., Carter, C., Silk, E., Stenger, V. A., Fissell, K., Goode, A. & Anderson, J. R. (2004). The change in brain activation patterns as children learn algebra equation solving. *Proceedings of the National Academy of Sciences USA 101*(15) 5686-5691.
- Quilodran, R., Rothe, M. & Procyk, E. (2008). Behavioral shifts and action validation in the anterior cingulate cortex. *Neuron 57*, 314-325.
- Raichle, M. E. & Snyder, A. Z. (2007). A default mode of brain function: A brief history of an evolving idea. *NeuroImage 37*, 1083-1099.
- Rakison, D. H. & Yermolayeva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Reviews: Cognitive Science 1* (in press).
- Reiss, S. (2004a). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology 8*(3), 179-193.

- Reiss, S. (2004b). The sixteen strivings for God. *Zygon* 39(2), 303-320.
- Rejendran, G. & Mitchell, P. (2006). Cognitive theories of autism. *Developmental Review* 27, 224-260.
- Ridderinkhof, K. R., van der Wildenberg, W. P. M., Segalowitz, S. J. & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition* 56, 129-140.
- Ring, H., Woodbury-Smith, M., Watson, P., Wheelwright, S. & Baron-Cohen, S. (2008). Clinical heterogeneity among people with high-functioning autism spectrum conditions: Evidence favoring a continuous severity gradient. *Behavioral and Brain Functions* 4(11) PMC2265729.
- Robinson, T. E. & Berridge, K. E. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews* 18, 247-291.
- Robinson, T. E. & Berridge, K. E. (2008). The incentive sensitization theory of addiction: Some current issues. *Philosophical Transactions of the Royal Society B363*, 3137-3146.
- Rolls, E. T. & Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: From affect to decision-making. *Progress in Neurobiology* 86, 216-244.
- Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition* 108, 771-780.
- Rushworth, M. F., Behrens, T. E., Rudebeck, P. H. & Walton, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behavior. *Trends in Cognitive Science* 11, 168-176.
- Sandkuhler, S. & Bhattacharya, J. (2008). Deconstructing insight: EEG correlates of insightful problem solving. *PLOS One* 3(1), e1459.
- Sarter, M., Gehring, W. J. & Kozac, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews* 51, 145-160.
- Saver, J. L. & Rabin, J. (1997). The neural substrates of religious experience. *Journal of Neuropsychiatry and Clinical Neurosciences* 9, 498-510.
- Saxe, R., Carey, S. & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology* 55, 87-124.
- Saxe, R., Tenenbaum, J. & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science* 16, 995-1001.
- Saxe, R., Tzelnic, T. & Carey, S. (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology* 43, 149-158.

- Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R. & Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Consciousness and Cognition* 17, 457-467.
- Scholl, B. & Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences* 4 (8) 299-309.
- Schubotz, R. & van Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observations and imagery. *The Journal of Neuroscience* 24 (24) 5467-5474.
- Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences* 3, 417-424.
- Shaw, M. P. (1999). On the role of affect in scientific discovery. In S. W. Russ (Ed) *Affect, Creative Experience and Psychological Adjustment* (pp. 147-162). London: Taylor & Francis.
- Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., Mintun, M. A., Wang, S., Coalson, R. S. & Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proceeding of the National Academy of Sciences USA* 106(6), 1942-1947.
- Smith, K. S., Mahler, S. V., Pecina, S. & Berridge, K. C. (2009). Hedonic hotspots: Generating sensory pleasure in the brain. In M. L. Kringelbach & K. E. Berridge (Eds.) *Pleasures of the Brain* (pp. 27-49). Oxford: Oxford University Press.
- Sobel, D., Yoachim, C., Gopnik, A., Meltzoff, A. & Blumenthal, E. (2007). The blicket within: Preschooler’s inferences about insides and causes. *Journal of Cognitive Development* 8(2) 159-182.
- Somerville, L., Heatherton, T. & Kelley, W. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience* 9, 1007-1008.
- Spikins, P. (2009). Autism, the integrations of 'difference' and the origins of modern human behavior. *Cambridge Archaeological Journal* 19(2), 179-201.
- Spikins, P. (2010). 'The world don't move to the beat of just one drum': Autism and the challenge of 'different minds' for the social sciences. Preprint, University of York.
- Subrahmanyam, K., Gelman, R. & LaFosse, A. (2002). Animate and other separately-movable objects. In E. Fordes & G. Humphreys (Eds.), *Category-specificity in brain and mind* (pp. 341-373). London, England: Psychology Press.
- Subramaniam, K., Kounios, J., Parrish, T. B. & Jung-Beeman, M. (2009). A brain mechanism for facilitation of insight by positive affect. *Journal of Cognitive Neuroscience* 21, 415-432.
- Vingerhoets, G., de Lange, F., Vandemaele, P. Deblaere, K. & Achten, E. (2002) Motor imagery in mental rotation: An fMRI study. *NeuroImage* 17, 1623-1633.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S. & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science* 7(3), 283-288.

Winkielman, P. & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology* 81(6), 989-1000.

Winkielman, P., Schwarz, N., Fazendeiro, T. A. & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds) *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*. Mahwah, NJ: Lawrence Erlbaum.

Winkielman, P., Halberstadt, J., Fazendeiro, T. & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science* 17(9), 799-806.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology, General* 136, 82-111.